Journal of Zankoy Sulaimani
Part-A- (Pure and Applied Sciences)
http://jzs.univsul.edu.iq

# Race Classification from Face Images Using Fast Fourier Transform and Discrete Cosine Transform

Hawkar O. Ahmed

*Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Kurdistan Region- Iraq*
*Department of Information Technology, University College of Goizha, Sulaimani, Kurdistan Region- Iraq*
**Corresponding author's e-mail: hawkar.omar@univsul.edu.iq*

**Article info**

Original: 28 April 2020
Revised: 25 June 2020
Accepted: 23 August 2020
Published online: 20 December 2020

***Key Words:***
FFT
DCT
Fusion
Knn

**Abstract**

Ethnicity identification and recognition is a key biometric technology with a wide range of applications related to homeland security, safety, access control, and automatic annotation. Ethnicity identification from face images is a process of gathering facial features of an individual face image compared to existing face images in the dataset to interpretation his/her ethnic class. In this paper, a propose method in multi-level fusion schema for ethnicity identification by using two global features; fast Fourier transform (FFT) and discrete cosine transform (DCT) on the pre-processed face image of size 128 * 128 in YCbCr color space. A dataset is consisting of 750 face image of three different ethnicities (Kurd 300, Oriental 300 and African 150). The query image feature is compared with a dataset image features using k – nearest neighbor classifier using City block distance for evaluating similarity measurement. The experimental result shows good accuracy and demonstrate the effectiveness of the combined features reached an accuracy rate 96.22% of classification.

## 1. Introduction

Within today's environment ethnicity identification and recognition are key biometric technology with a wide range of applications related to homeland security, safety, access control, and automatic annotation, etc... It is generally understood that ethnic group is the biological unit of human classification, with genetic physical features making up the group classification. The most common features include head shape, face shape, skin color. Ethnic groups are based on the most critical characteristics in the human faces such as color, texture and shape for all attributes like eyes, nose, mouse, lips and checks. With uncontrollable conditions such as lighting, eye glasses and expressions, it is difficult for researcher to use an algorithm for extracting features from human faces, then search for similar facial images from a given dataset, and finally determine the ethnic group of the query face image.

## 2. Literature Survey

Ethnicity identification and classification using face images is a relatively new topic in computer vision [1]. The algorithms used to represent the face images for ethnicity identification and classifications are few and limited. Early works dealt with two-class situations (Asian vs. non-Asian) [2, 3]. Jain in [4] used a linear discriminant analysis (LDA) technology for facial features extraction to classify the Asians and non-Asians and the K-nearest neighbor (KNN) method has been used (add) for classification. Hosoi and Takikawa [5] used Gabor Wavelets Transformation to extract key facial features and used support vector machines for ethnicity classification. More recently, Salah et al [6] used local binary patterns as local features and Haar wavelet transform as global features then concatenate local and global features to produce fusions system with a simple k-nearest neighbor classifier for ethnicity identification with a good accuracy. Faraidoon and Aree

[7], describe a hybrid strategy that uses Haar wavelet and discrete cosine transform, the experiment results show good level of accuracy for ethnicity identification. More recently study, Hongbo Du [8] presents an approach through fusion of color feature, by using Hue-Saturation-Value (HSV) and texture feature by using Local Binary Pattern (LBP) with a K-NN and Support Vector Machine (SVM) classifier for three ethnic groups, the result showed good level of accuracy. In this approach, present a new ethnicity estimation technology based on a new multi-level fusion scheme that attempts to combine both two global feature extractors DCT and FFT from the image and apply KNN to estimation, with the use of City block for determine similarity measurement, on collecting of 750 images.

## 3. Feature Extraction Methods

### 3.1 *Discrete Cosine Transform*

Discrete Cosine Transform (DCT) is a popular transform method in image processing or in signal processing that is used to transform an image from the spatial domain to the frequency domain it is derived from the Discrete Fourier Transform (DFT). The DCT is applied on 2-D image for extracting features. DCT was applied to the process of face recognition [10] and used DCT for the processing of finger print identification.

The general equation for a 2D (N by M image) DCT is defined by the following equation:

$$F[x,y] = \frac{2}{\sqrt{XY}} \mathrm{I(x)I(y)} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} f[i,j] cos\frac{(2i+1)x\pi}{2X}\, cos\frac{(2j+1)y\pi}{2Y} \qquad \ldots\ldots(1)$$

$$I(x)I(y) = \begin{cases} 1/\sqrt{2} & x,y=0 \\ 1 & otherwise \end{cases}$$

Where

- The size input image is N by N matrix.
- f(i,j) is the intensity of the pixel in row x and column y.
- F(x,y) is the DCT coefficient in row x and column y of the DCT matrix.

In this work, after converting the original image from RGB color space to YCbCr color space DCT operator are implemented on Y channel as follows:

- Divided the Y channel into (8×8 pixels) sub blocks.
- Each (8 bit) pixels have levels from 0 to 255.
- The DCT is performed over each sub block independently.
- Implementing the zigzag scanning function on each sub block.
- The output array of DCT coefficients on each block contains 64-element DCT transform (1 DC coefficient and 63 AC coefficients).
- The lower right values represent higher frequencies and are of less visual importance.
- The upper left values represent low frequency and contains most of information.

### 3.2 *Fast Fourier Transform*

Fast Fourier Transform (FFT) is only a fast process for computing the Discrete Fourier Transform (DFT). DFT is used to convert one function in the spatial domain into the frequency domain. DFT is widely used in signal processing applications such as correlation analysis, spectrum analysis and linear filtering that includes less efficient algorithms and more computational time resulting. FFT is computationally faster than DFT by decomposing sequence into a small sequence until signal point sequences are got [11]. The FFT of a 2D image are given by the following equations.

$$F(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) e^{-j2\pi(x\frac{m}{M}+y\frac{n}{N})} \qquad \ldots\ldots\ldots\ldots(2)$$

$$f(m,n) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(x,y) e^{j2\pi(x\frac{m}{M}+y\frac{n}{N})} \qquad \ldots\ldots\ldots\ldots(3)$$

Where:

- f(m,n) is the pixel at coordinates (m,n),
- F(x,y) is the value of the image in the frequency domain corresponding to the coordinates x and y, and M and N are the dimensions of the image.

As illustrated in (Fig. 1) when FFT is applied on an image, the black image will appear because of complex calculation but in the next step the DC-value will appear in the center of the image in the form of largest component of the image. The result of FFT shows that the image contains components of all frequencies; high frequency contains less information than the lower ones. In the Fourier image two domain directions are presented: one passing horizontally and one vertically through the center.

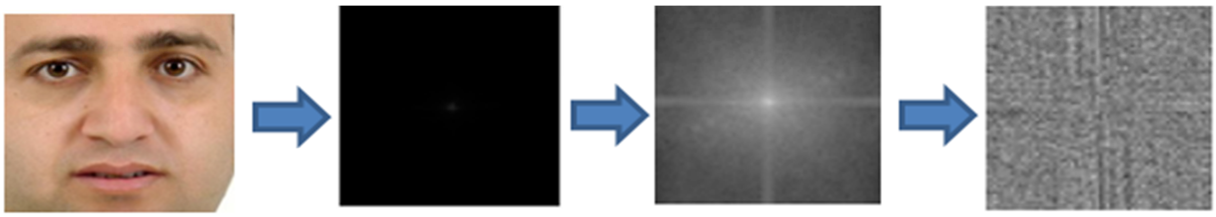

Figure-1: The phases of the Fourier transform of the image

## 4. The Framework

In this paper, a new fusion schema is proposed that concatenate both (DCT and FFT) global feature extractors. The framework of the proposed schema has been illustrated as a block diagram in Fig 1. It contains several steps such as preprocessing, converting color space, feature extraction, feature fusion and classification.

The Framework is described in the following steps:

1. In Preprocessing step: the frontal images are cropped to the fixed size 128*128.
2. Converting color space: in this step the original RGB color image is converted in to YCbCr color space. In this color space, only Y channel is used because it contains the texture of an image
3. Feature extraction: two types of global feature extractor are implemented FFT and DCT separately on each 8 * 8 blocks.
4. Feature fusion: In the fusion step, the feature-level fusion is simply the concatenation of the DC coefficient in FFT feature extractor with the different coefficient in two dimensional DCT as explained in section 3.1 and the values of the two global feature extractors are normalized using the division by norm normalization technique. Finally, a feature vector for the whole frontal image is produced.
5. Classification: to determine the class of the query image, the KNN classifier is used and select 5 nearest face images according to the City block distance between the feature vectors of query image and training images and, the City block distance is defined in equation (4)

$$D(x,y) = \sum_{i=0}^{k-1} |x_i - y_i| \quad \dots\dots\dots (4)$$

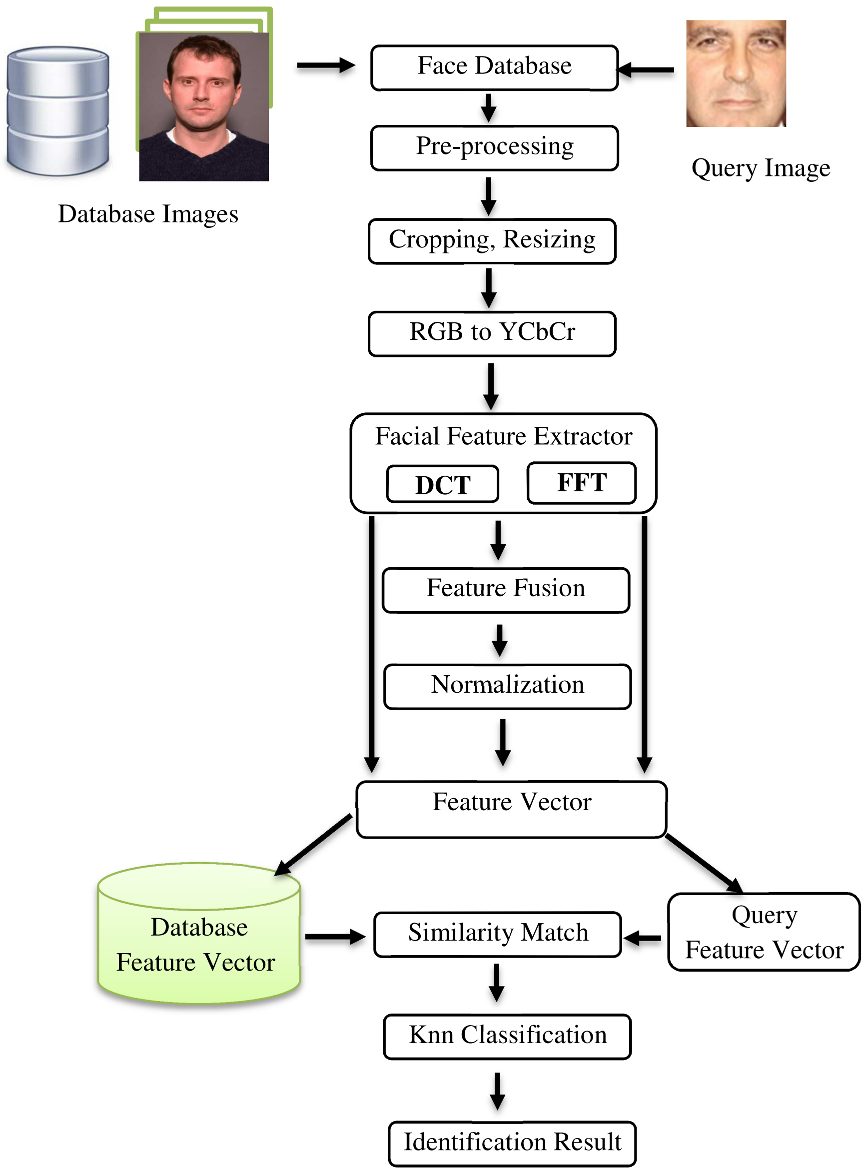


Figure-1: The Framework of the Proposed Method

## 5. Experiments and results

### 5.1 *Dataset*

At this time, there is not a standard dataset available for ethnicity identification. therefore, 750 frontal images from five different face dataset s are collected such as FERET dataset [12], the YALE face dataset [13], the PIE face dataset [14], the GUFD face dataset [15] and a snapshot dataset. Which consist 200 frontal images from FERET dataset, 80 frontal images from PIE dataset, 70 frontal images from GUFD dataset, 50 images from YALE dataset and the snapshot frontal image dataset contains 300 Kurd and 50 African snapshot frontal images. This dataset is divided in to three ethnic classes (European, Kurd and African); the numbers of frontal images for each ethnicity are 300, 300 and 150 respectively.

### *5.2 Results and Discussion*

In this section, the performance of the proposed method for ethnicity identification from the facial images are showed by using a collection of 750 RGB face images in different standard dataset images as illustrated in Section 5.1. These images are cropped with a fixed size of 128 x 128 pixels, and then converted from RGB color space to YCbCr. In experimental results, two different algorithms (DCT and FFT) for extracting global features are applied from each face images in the dataset separately for identification process, then implementing fusion strategy that combines DCT coefficient and FFT coefficient to produce one feature vector for representing the image. The values of the feature vector are normalized by using the division by norm normalization technique. As numerical feature vectors are produced by algebraic features, the KNN classifier [16, 17] is used to make prediction to the test. In this study finds 5 nearest images are retrieved according to the City block distance between the feature vector of the training images and testing image.

In the feature extraction process, the global features are extracted by applying DCT over all face images, then different length of DCT feature vectors from upper left corner are selected because it contains most of information such as (4, 9, 16, 25 and 36) for block sizes 2×2, 3 × 3, 4 × 4, 5 × 5 and 6 ×6 respectively. Table 1 shows the test results for the average accuracy of identification at different DCT block size on the dataset for three ethnicity group (Kurd, Oriental and African).

Table-1: Average accuracy using different DCT block size for three Ethnic group

| **DCT block size** | **Ethnicity Groups** | | | **Average Accuracy** |
|---|---|---|---|---|
| | Kurd | Oriental | African | |
| **2×2** | 99.67 | 98.33 | 92.67 | 96.89 |
| **3 × 3** | 99.33 | 99.33 | 96.67 | 98.44 |
| **4 × 4** | 99 | 99.67 | 96 | 98.22 |
| **5 × 5** | 99.33 | 100 | 96.67 | 98.67 |
| **6 ×6** | 99.33 | 98.33 | 87.33 | 95 |

The table illustrates that the classification accuracies for 5 × 5 block size has better accuracy than the other block size by achieving 98.67% accuracy.

The Second method for extracting features is FFT, in this case DC-value as illustrated in section 3.2 are extracted for some insight, a confusion matrix is presented by using the cross validation for evaluation of accuracy with implementing the leave-one-out strategy. Table 2 presented the result of three ethnicity groups (Kurd, Oriental and African), the accuracy rates are high and better in Kurd group and Oriental group which are 99.33%, 98.33% respectively but in African group the result is worse with average accuracy of 87.33%.

Table-2: A Confusion Matrix Representing Accuracy for Three Ethnic Group using FFT

| | | **Predicted Ethnicity** | | |
|---|---|---|---|---|
| | | **Kurd** | **Oriental** | **African** |
| **Actual Ethnicity** | Kurd | 99.33 | 0.67 | 0 |
| | Oriental | 0.33 | 98.33 | 1.33 |
| | African | 0.67 | 12 | 87.33 |

Based on preliminary investigations reported in the previous section, a fusion scheme that combines DCT coefficients and FFT coefficient to produce one feature vector for representing the image are proposed, then for evaluating the performance of the proposed fusion schema, the DC-value of FFT feature extractor with different block size of DCT feature extractor including (4, 9, 16, 25 and 36) for block sizes 2×2, 3 × 3, 4 × 4, 5 × 5 and 6 ×6 respectively are combined. Then the values of the new feature vectors are normalized by using the division by norm normalization technique, after that the KNN classifier is are used to make predictions to the test. In this study finds 5 nearest images are retrieved according to the City block distance between the feature vector of the training images and testing image. Table 3 shows 10 features achieved the best overall performance, achieving the average accuracy of 96.22%. for global features. 17 features produced a poorer performance with average accuracy of 91.11%.

Table 3: Ethnicity Identification Accuracy for the Fusion Method

| No. of Features | DCT block size | Ethnicity Groups | | | Average Accuracy |
|---|---|---|---|---|---|
| | | Kurd | Oriental | African | |
| 5 | 2×2 | 96.00 | 96.33 | 90.00 | 94.11 |
| 10 | 3 × 3 | 96.67 | 97.33 | 94.67 | 96.22 |
| 17 | 4 × 4 | 88.33 | 99 | 86 | 91.11 |
| 26 | 5 × 5 | 93.33 | 99.67 | 90.67 | 94.56 |
| 37 | 6 ×6 | 91.67 | 99 | 87.33 | 92.67 |

In order to gain better understanding towards the effects of the DCT features and FFT feature on the three ethnicities, a confusion matrix is presented in table 4 for the 10 feature (DC-value of FFT and 3 × 3 block size of DCT) Across ethnicities, the accuracy rates are similar, slightly better in the Kurd and Oriental classes and slightly worse in the African class.

Table 4: Confusion Matrix for DC-value of FFT and 3 × 3 block size of DCT

| | | Predicted Ethnicity | | |
|---|---|---|---|---|
| | | **Kurd** | **Oriental** | **African** |
| **Actual Ethnicity** | **Kurd** | 96.67 | 3.33 | 0 |
| | **Oriental** | 2 | 97.33 | 0.67 |
| | **African** | 1.33 | 4 | 94.67 |

## *5.3 Comparison to Existing Results*

Comparison of the proposed technique with similar recent works is difficult. This is due to non-availability of the common frontal facial image database which takes into account standard environment conditions. Comparisons were done with some previous studies. Table 5 shows the comparison of the proposed methods results with that of [8] [18] [19] and shows that the proposed method has superior accuracy of identification than the other methods with 10 selected features. It is inappropriate to compare the proposed method with [2] and [3] as these deal with binary class (Asian vs. Non-Asian) situation only. The introduction of a third class as in the proposed methods would affect accuracy.

Table 5: Comparison between proposed scheme with recent works.

| Ref. No. | Feature Selected | Kurd | Oriental | African | Average |
|---|---|---|---|---|---|
| [8] | HSV+LBP | 94.78 | 95.36 | 96.23 | 95.46 |
| [18] | GWT +Retina Sampling | 93.1 | 96.33 | 94.3 | 94.56 |
| [19] | LBP | 96.60 | 93.60 | 93.80 | 94.67 |
| Proposed Scheme | FFT + DCT | 96.67 | 97.33 | 94.67 | 96.22 |

## 6. Conclusion

In this paper, we have presented a fusion scheme for ethnicity identification based on texture feature. The facial images are first subjected to some image preprocessing algorithm to help the process of identification. A special database containing 750 images is created for three different ethnic groups (Kurd, Oriental and African). The fusion scheme fuses the DC-value of FFT with combined different block size of DCT including (4, 9, 16, 25 and 36) for block sizes 2×2, 3 × 3, 4 × 4, 5 × 5 and 6 ×6 respectively into a single bin feature vector, and then applies KNN classifiers to the combined feature vector. The experiment results show high degree of accuracy was achieved by 10 features, and the recognition rate was 96.22 %.

## References

[1] Ghulam M., Muhammad H. and Fatmah A. *"Race Classification From Face Images Using Local Descriptors"*. International Journal on Artificial Intelligence Tools. Vol. 21, No. 5. (2012).

[2] Lu, X. and Jain, A. K. *"Ethnicity identification from face images"*. Proc. SPIE International Symposium on Defense and Security: Biometric Technology for Human Identification, Orlando, Florida, April. pp. 114-123. (2004).

[3] Manesh, F. S., Ghahramani, M., and Tan, Y. *"Facial part displacement effect on template-based gender and ethnicity classification"*. Control Automation Robotics & Vision (ICARCV), Proc. 11th International Conference on Control, Automation, Robotics and Vision, Singapore, December, pp. 1644-1649. (2010).

[4] Xiaoguang Lu, Anil K. Jain. *"Ethnicity Identfication from Face Images [J]"*. Bull.Inst. Math. Acad. Sinica. Vol. 33, pp. 77-87. (2005).

[5] Satoshi Hosoi, Erina Takikawa, Masato Kawade. *"Ethnicity Estimation with Facial Image [C]"*. Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 10-16. (2004).

[6] Salah, S. H., Du, H. and Al-Jawad, N. *"Fusing local binary patterns with wavelet features for ethnicity identification"*. Proc. ICSIP 2013: International Journal of Computer, Information Science and Engineering. Vol. 7, No. 7, pp. 330-336. (2013).

[7] Faraidoon H. & Aree A. *"Hybrid Wavelet and Discrete Cosine Transform Methods for Ethnicity Identification"*. JZS-A. Vol. 17, No. 1. (2015).

[8] Hongbo Du; Sheerko H. Salah; Hawkar O. Ahmed. *"A color and texture based multi-level fusion scheme for ethnicity identification"*. Published in Proceedings Volume 9120: Mobile Multimedia/Image Processing, Security, and Applications (June 2014).

[9] Ziad M. Hafed and Martin D. Levine. *"Face recognition using the discrete cosine transform*". Vol. 43, pp. 167–188. (2001).

[10] Ritu K. and Susmita M. *"Fingerprint Based Gender Identification Using Frequency Domain Analysis"*. IJAET. March (2012).

[11] Y. Rangaswamy; K. B. Raja; K.R. Venugopal. *"FRDF: Face Recognition using of DTCWT and FFT Features"*. ScienceDirect, Procidia Computer Science. Vol. 54, pp. 809-817. (2015).

[12] P. Phillips, H. Moon, S. A.Rizvi, and P. J. Rauss. *"The feret evaluation methodology for face-recognition algorithms*". IEEE Trans. PAMI. Vol. 22, No. 10, pp. 1090–1104.

[13] Burton, A., White, D., and McNeill, A. *"The Glasgow face matching test. Behavior research methods*". Vol. 42, No. 1, pp. 286-291. (2010).

[14] T. Sim, S. Baker, and M. Bsat. *"The cmu pose, illumination, and expression (pie) dataset*". Pp. 53–58. (2002).

[15] Milborrow, S., Morkel, J. and Nicolls, F. *"The MUCT landmarked face dataset*". Pattern Recognition Association of South Africa. (2010).

[16] Teknomo K. *"What is K Nearest Neighbors Algorithm"*. htp://people.revoledu.com/kardi/tutorial/KNN/Contents.htm.

[17] D'Amato C,Malerba D,Esposito F,et al. *"Extending the K-Nearest Neighbour classification algorithm to symbolic objects[C]* ". Convegno Scientifico Intermedio SIS. Universit -àdegli Studi di Napoli"FedericoĊ, 9-11 Giugno (2003).

[18] S. Hosoi, E. Takikawa, and M. Kawade, *"Ethnicity estimation with facial images"*. In Proc. 6th IEEE Int. Conf, on Automatic Face and Gesture Recognition (AFGR). pp. 195–200. (2004).

[19] Hawkar O. Ahmed, Mahdi M. Younis and Shakhawan H. Wady. *"LBP Variants as Texture Descriptors for Ethnicity Identification"*. JZS-A. Vol. 18. No.3. (2016).